NYSERDA Artificial Intelligence Tools and Process

Dana Nilsson

NEW YORK STATE | NYSERDA

Background

In March 2025, New York State established comprehensive AI guidelines for state agencies, based on the NIST AI Framework

https://its.ny.gov/system/files/documents/2026/01/nys-p24-001-acceptable-use-of-ai.pdf

- Key criteria include: Human oversight, Fairness & Equity, Transparency, Risk Assessment, and Privacy
- All use cases must be approved by internal NYSERDA AI Governance to ensure compliance with criteria

Challenges include:

- Designing test scenarios that contain a sufficient number of human checks that simultaneously balance project costs, project timeline, and resource allocation while also satisfying AI Governance criteria
- Selection and/or development of tools that could run entirely on NYSERDA servers, or contractor-owned, restricted access on-site servers without data leakage through cloud AI providers. This requires NYSERDA or contractor to have on premise GPUs and teams able to develop and maintain the technical infrastructure around the AI model
- All content created by the AI model must be clearly identified and stored separately from other data
- Any reports that contain information generated by AI must disclose such, and be reviewed in advance of publication by AI Governance

Use of AI at NYSERDA

Likely Permitted AI Systems	Likely Prohibited AI Systems
• **AI Systems that generate code; provided human operators have appropriate training, oversee the entire process, and review the final code before implementation** • **AI Systems for content editing; requires prior approval, qualified human supervision and safety guardrails in place to preserve content integrity and information security** • **Secure internal LLM's**	• Applications with LLM's that train on non-public NYSERDA data • AI Systems that log keystrokes/activities • AI Systems that listen to or record internal meetings without prior approval

Current use cases for evaluation

- Use cases where NYSERDA has found success include:
  - Creating data sets from information sources (internal or public) that are unstructured. Data and reports needed for evaluation often contain non-standardized text and descriptors, images, and tables
  - Aggregating information from company websites in order to better characterize the market
  - Efficient and timely processing of information needed to inform sampling, interviews and proxy estimates
  - Verification of energy efficiency measures assigned zero savings
  - Efficient resolution of missing or unreported savings

AI tool developed by DNV on behalf of NYSERDA extracted usable data from 345 sites (86%) of single-project sites in the Energy Management Technology evaluation* period (2017-2024). Web data gathering also employed AI and gathered data from nearly 8,400 pages covering 211 non-participating vendors

*https://www.nyserda.ny.gov/-/media/Project/Nyserda/Files/Publications/PPSER/Program-Evaluation/2025-11-RTEM-Commercial-Industrial-Direct-and-Indirect-Benefits-Impact-Study.pdf

The Future and next steps

- Development of methods and models that can be straightforwardly reviewed for critical AI criteria without starting from scratch on internal approvals
- Removing barriers while ensuring security and reliability
- Methods/models that can easily be leveraged for other data management tasks on data already approved for AI
- Development of a path for AI analysis, data management, and storage on non-site servers (either by NYSERDA or consultant)

Power TakeOff

IEPEC Webinar: AI in EM&V

How Should Evaluators use AI?

Lucas Born, Senior Manager, EM&V

April 2026

Supercharging M&V with AI

Tier 1 — Arming M&V Teams
(background & process)

LLMs as a learning accelerator

Getting up to speed faster on new concepts

Automating data cleaning & NRE detection

AI coding agents / Scripting access

- Moving the industry off spreadsheets
- Expanding Horizons: ML & 3rd Party Software

Tier 2 — Reimagining Calculation
(the novel frontier)

ML (Machine Learning) niche variables vs traditional linear regression.

Non-linear models outperform standard IPMVP options in accuracy.

The Payoff:
Precise savings → Credible DSM → DSM as a real grid resource

The Transparency Tightrope

Core Tension

The friction between innovation and oversight.

- Managing PII and data governance when using third-party AI tools
- Blackbox ML power vs. regulatory need for auditability and replication
- Current M&V frameworks weren't designed for non-linear, probabilistic models

Who Blinks First?

Evaluators, regulators, or implementers?

- Evaluators as trusted third parties may be best positioned as first movers
- Proving AI models are accurate *and* auditable gives regulators cover to approve
- **Open Question:** Does "replicable" need to be redefined for an AI-augmented world?

EcoMetricx

IEPEC WEBINAR · APRIL 2026

How Should Evaluators Use AI?

A practitioner's view on where AI moves the needle for energy program evaluation — and where evaluators must keep their hands on the wheel.

PRESENTED BY

Matthew Harding

CEO, EcoMetricx · harding@ecometricx.com

01 WHY NOW

Why AI and Evaluation?

Evaluation work is data-rich, repetitive, and methodologically demanding. AI shifts evaluators from production to insight.

Volume & Velocity

AMI, weather, EnergyStar, CRM — large-scale signals that legacy workflows can't keep pace with.

Methodological Rigor

LLMs accelerate UMP-aligned model selection, diagnostics, and documentation without skipping the science.

Transparency on Demand

Auto-generated rationale, lineage, and narrative reporting that regulators and clients can actually read.

From Lookback to Foresight

Continuous causal estimates feed forecasts, targeting, and program design — not just compliance.

02 THE SHIFT

Move utility EM&V from annual retrospectives to live, automated, decision-grade insight.

FROM · ANNUAL EM&V

- Periodic, retrospective reports
- Frozen data extracts and manual QC
- Findings arrive after decisions are made

TO · CONTINUOUS, EMBEDDED

- **Always-on causal estimates per program**
- **Automated pipelines on live AMI streams**
- **Insights inform targeting in real time**

Live Pipelines
AMI → causal model

Auto-Refresh
Monthly impact updates

Faster Insight
Targeting & forecasting

Audit-Ready
UMP-aligned every run

Evaluation as a service — not an annual deliverable.

03 THE GUARDRAILS

Data Security & Governance

Customer energy data is sensitive. AI in EM&V only works if security and governance lead the architecture.

Single-Tenant Deployment

Models and pipelines run inside the utility's own cloud — no multi-tenant data pooling.

Encryption & Key Control

S3/KMS encryption, customer-managed keys, and TLS in transit by default.

Least-Privilege IAM

Role-scoped access for every job, model, and human reviewer.

Privacy by Design

PII minimization, customer aggregation thresholds, and policy gates on model inputs.

Immutable Lineage

Every run produces audit artifacts: data versions, model rationale, diagnostics, code.

Human Oversight

Critical findings escalate to reviewers before they influence program decisions.

Trustworthy AI in EM&V starts with trustworthy data plumbing.